

Large-scale use of an objective, structured clinical examination for licensing family physicians

Paul Grand'Maison, MD, MCISc, FCFP; Joëlle Lescop, MD, MPH, FRCPC, FCFP; Paul Rainsberry, PhD; Carlos A. Brailovsky, MD, MA

Since 1988 in Quebec the completion of a residency training program in family medicine or a specialty and of a comprehensive examination has been necessary to obtain a licence. An objective, structured clinical examination (OSCE) was designed by the Corporation professionnelle des médecins du Québec and Quebec's four medical schools to evaluate the clinical competence of newly trained family physicians. The certification examination of the College of Family Physicians of Canada was added to the OSCE. More than 500 candidates have been assessed: 262 in the spring of 1990, 42 in the fall of 1990 and 235 in the spring of 1991. The spring session occurs in four centres, three offering it in French and one in English, and the fall session takes place in one bilingual centre. In each centre 25 standardized patients and 25 examiners are required on each day of the 2-day OSCE. The scores obtained by the candidates who completed the OSCE in the first three sessions showed a normal distribution. No more than 5% failed the OSCE, the pass level having been set at two standard deviations below the mean. Equivalence was shown among the OSCE tracks, and reliability coefficients of 0.644, 0.723 and 0.736 were obtained for the three sessions respectively. The overall success rate for the licensing examination was 92%. The integration of such a large-scale OSCE into a licensing examination and the results obtained show that assessment of clinical competence for licensing purposes is feasible. The Quebec experience may help other organizations that are developing OSCEs for summative purposes.

Depuis 1988 au Québec, l'obtention du permis d'exercice exige de compléter un programme de résidence en médecine de famille ou en spécialité et de réussir l'examen terminal correspondant. La Corporation professionnelle des médecins du Québec et les quatre facultés de médecine du Québec ont développé un examen clinique objectif et structuré (ECOS) pour évaluer la compétence clinique des médecins de famille nouvellement formés. À cet examen s'ajoute l'examen de certification du Collège des médecins de famille du Canada. Déjà, plus de 500 candidats ont été évalués : 262 au printemps 1990, 42 à l'automne 1990 et 235 au printemps 1991. La session du printemps a lieu dans quatre centres d'examen, dont trois en français et un en anglais. La session d'automne a lieu dans un centre bilingue. Dans chaque centre, 25 patients standardisés et 25 examinateurs sont nécessaires chaque jour au déroulement de l'ECOS qui dure 2 jours. L'analyse des résultats des candidats qui ont participé aux trois

Dr. Grand'Maison is professor and head, département de médecine de famille, Université de Sherbrooke, Sherbrooke, Que.; Dr. Lescop is director, Service des études médicales, Corporation professionnelle des médecins du Québec, Montreal, Que.; Dr. Rainsberry is director of education, College of Family Physicians of Canada, Mississauga, Ont.; and Dr. Brailovsky is professor and director, Centre d'évaluation en sciences de la santé de l'Université Laval, Faculté de médecine, Université Laval, Sainte-Foy, Que.

Reprint requests to: Dr. Paul Grand'Maison, Département de médecine de famille, Faculté de médecine, Université de Sherbrooke, 3001, 12^e ave nord, Sherbrooke, QC J1H 5N4

premières sessions de l'ECOS a montré une distribution normale de leurs scores. Avec une note de passage fixée à deux écarts types sous la moyenne, moins de 5 % des candidats ont échoué à l'ECOS. L'analyse des différents circuits ECOS a démontré qu'ils étaient équivalents. Des coefficients de fidélité de 0,644, 0,723 et 0,736 ont été obtenus respectivement pour les trois sessions. Le taux global de réussite à l'examen du permis a été de 92 %. L'expérience démontre qu'il est possible d'évaluer la compétence clinique en intégrant un ECOS d'envergure à un examen du permis. Cette expérience québécoise devrait servir à tous ceux qui projettent d'utiliser des ECOS à des fins d'évaluation sommative.

For medical licensing authorities accountability means granting a licence to competent physicians and assuring the public that those physicians remain competent throughout their professional lives. For medical schools accountability means training physicians to have up-to-date knowledge, skills and attitudes. To assure accountability the Corporation professionnelle des médecins du Québec (CPMQ) and the four Quebec medical schools proposed stringent licensure regulations, which were adopted by the Quebec government in 1987.¹ All medical students graduating in 1988 and thereafter and wishing to obtain a licence to practise in Quebec have had to complete an accredited program in family medicine (2 years) or in one of the specialties (4 or 5 years) and pass a terminal examination assessing clinical competence in their field.

The licensing examination for the first cohort of family medicine residents had to be in place by May 1990. A policy board of eight people, four nominated by the CPMQ and one by each of the four medical schools, was responsible for designing the terminal examination to assess cognitive, clinical and affective skills. An objective, structured clinical examination (OSCE)² had to be developed to measure clinical competence. The certification examination of the College of Family Physicians of Canada (CFPC),³ which has been used for the last 20 years, would also be part of the licensing examination to complement the OSCE. This paper presents the development, implementation and results of the first three sessions of the OSCE.

OSCE

The past two decades have seen an increasing interest in the assessment of future physicians' clinical competence and skills that were not assessed adequately by traditional means. This has led to a new emphasis on the use of performance-based tests such as the OSCE in which the tasks presented to examinees are more representative of those faced in real clinical situations. Medical schools worldwide now use these tests.^{4,5}

The OSCE was introduced to assess components of medical competence. Examinees rotate through a

circuit of clinical stations in which components of their clinical skills are assessed in a planned, structured and objective way through direct observation.² The examinee's performance is scored on checklists or rating forms tailored to the content of each station. The clinical tasks on which the assessment is based may include history-taking, performing some portion of a physical examination, performing a procedure, requesting diagnostic tests, interpreting test results, diagnosing, planning management and treatment or communicating with a patient.

To test the clinical skills standardized patients are used.^{6,7} A number of studies, reviewed by van der Vleuten and Swanson,⁸ have shown that simulated patients can play the role of patients in a standardized and consistent way and that exams involving such people can obtain reliability coefficients of 0.41 to 0.88. Although OSCEs using standardized patients have proven to be effective assessment tools over the years,⁴⁻⁷ Quebec has been the first to use an OSCE on a large scale for such an important issue as licensing.

Quebec experience

The Quebec OSCE was developed to measure the clinical skills necessary to practise as an independent family physician. An examination committee of 10 family physicians from urban, rural and teaching practices developed clinical cases according to the following principles.

- The cases must be designed on the basis of important problems seen in family medicine (e.g., low back pain, sexually transmitted diseases and chest pain). The age and sex distribution must be close to that of real practice, and all body systems must be represented.

- The clinical settings must be ones in which family physicians practise (75% of the cases would be presented in an office setting).

- Assessment and observation of clinical skills must be maximized. In 75% of the cases the family physician examiners would use direct observation to assess the candidates' competence.

The spring 1990 OSCE included 26 cases distributed over 40 stations. Each station took 7 minutes to complete. The candidates were required to go

through one OSCE track of 2½ hours' duration on each of 2 consecutive days. Each track consisted of 13 cases and 20 stations. Standardized patients were used in 23 cases. Depending on the station the candidates were asked to take a patient's history, perform parts of a physical examination, complete a specific technical act or discuss the problem, the diagnosis and the treatment with the patient. Most stations assessed many of these tasks. Direct observation and scoring with the use of checklists was done by raters, all of whom were family physicians. Fourteen of the stations involved completion of a questionnaire on information the candidate obtained at the previous clinical station. The questionnaires focused on issues such as data interpretation, diagnosis, investigation and treatment. All of the questions were of the short-answer or short-menu type.

Similar OSCEs were developed for the fall 1990 and spring 1991 examinations. More than 50% of the cases at each session were new. A few 14-minute clinical stations were added because of their suitability for more complex cases.

A short multiple-choice questionnaire of 50 items was also developed to assess knowledge of the legal, ethical and organizational aspects of medical practice in Quebec. This part of the examination was identified by the acronym ALDO-Quebec (les aspects législatifs, déontologiques et organisationnels de la pratique médicale au Québec).⁹

CFPC examination

The certification examination of the CFPC is designed to assess the knowledge and skills of trainees headed for careers in family medicine.³ It evaluates how well candidates meet the college's objectives: to be able to define health problems, show good affective skills, manage health problems effectively, integrate health maintenance activities, show professional responsibility, use knowledge of the family and understand research methods.

The examination uses multiple-choice questions (220 items), short-answer management problems (20 clinical cases) and oral examinations. Each of the five simulated office oral exams simulates a 15-minute office visit in which the examiners play the roles of patients. Items from each instrument are assigned to specific educational objectives. Candidates pass or fail on the basis of how well they meet the objectives rather than how well they do on each individual instrument.

Implementing the licensing examination

Le Centre d'évaluation en sciences de la santé de l'Université Laval is responsible for implementing the OSCE. A chief patient instructor trains one

patient instructor for each of the four exam centres. This person then trains the simulated patients during the weeks preceding the examination. A chief coordinator is responsible for training four coordinators, who in turn train the examiners the day before the OSCE is administered. Scenarios, videotapes depicting the cases, and the scoring sheets are used to train and standardize patients and observers. The CFPC is responsible for administering the certification examination.

The entire Quebec licensing examination lasts 3 days, the first day being devoted to the written examination and the following 2 days to the OSCEs and the simulated office oral exams.

The spring examination, which involves about 250 candidates, takes place simultaneously at all four centres: in French at Université Laval, Université de Montréal and Université de Sherbrooke and in English at McGill University. At each centre three or four identical and parallel tracks are used. Candidates are scheduled in such a way as to assure confidentiality of the tracks. A total of 50 to 80 candidates are assessed at each centre with the use of 25 standardized patients and 25 OSCE examiners each day.

The fall examination takes place at only one bilingual centre and involves about 50 candidates.

Results

OSCE

In the spring of 1991 the scores for the OSCE showed a normal distribution (Fig. 1). The mean score was 71.3% (standard deviation 4.8%). The

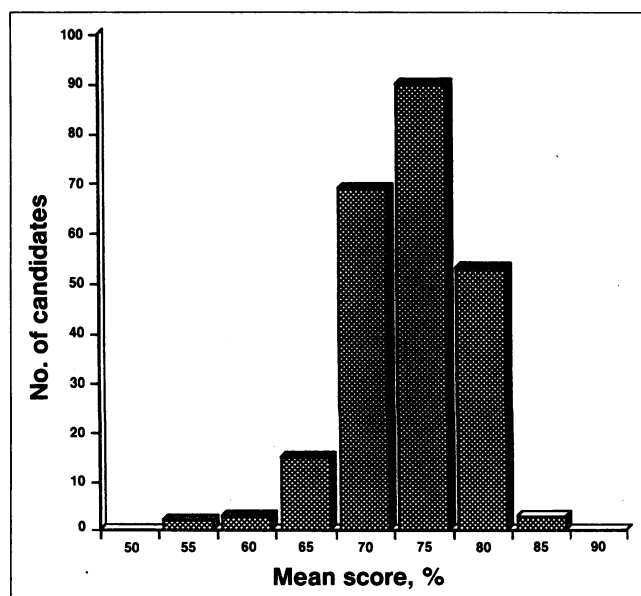


Fig. 1: Distribution of OSCE scores for 235 candidates in the spring of 1991.

negative skewness of the score distribution was expected since the exam occurred at the end of a specific training program. The equivalence of the parallel tracks used in the OSCE (14 in the spring and 3 in the fall) was assessed with the use of one-way factorial analysis of variance.¹⁰ The Barlett test¹¹ was used to verify the homogeneity of group variances. Once the homogeneity was confirmed a Tukey-Kramer Honestly Significant Differences test,¹² adjusted for unequal numbers of replications,¹³ was applied for the post-hoc comparison of the means. The results confirmed that the trainees in each track behaved like samples from the same population.

The candidates' overall results were derived from the sum of their scores in each of the cases, all cases being weighted equally. The results for the first three examinations are shown in Table 1.

The distribution of the scores at the four centres varied from case to case because of the relative difficulty and structure of each case, as illustrated in the spring 1991 example (Fig. 2). This justifies the use of a large number of cases to assure content validity and reliability. The reliability was calculated with the use of the Testat module for Systat,¹⁴ each case being considered as one item. Using Cronbach's α internal consistency analysis¹⁵ we calculated the reliability coefficients of the three examinations to be 0.644, 0.723 and 0.736 (Table 1). In addition, Fig. 2 illustrates the lack of difference between the four centres in overall scores for each case.

The pass level was set for the first three examinations at two standard deviations below the mean. A normative approach that would fail a small proportion of candidates was chosen deliberately for

Table 1: Statistical data for the first three sessions of an objective, structured clinical examination used in Quebec for licensing physicians in family medicine

Variable	Spring 1990	Fall 1990	Spring 1991
No. of candidates	262	42	235
Mean score (and standard deviation), %	68.8 (4.9)	70.1 (4.9)	71.3 (4.8)
Standard error of measurement	2.999	2.543	2.466
Reliability coefficient	0.644	0.723	0.736
Success rate, %	97	100	95

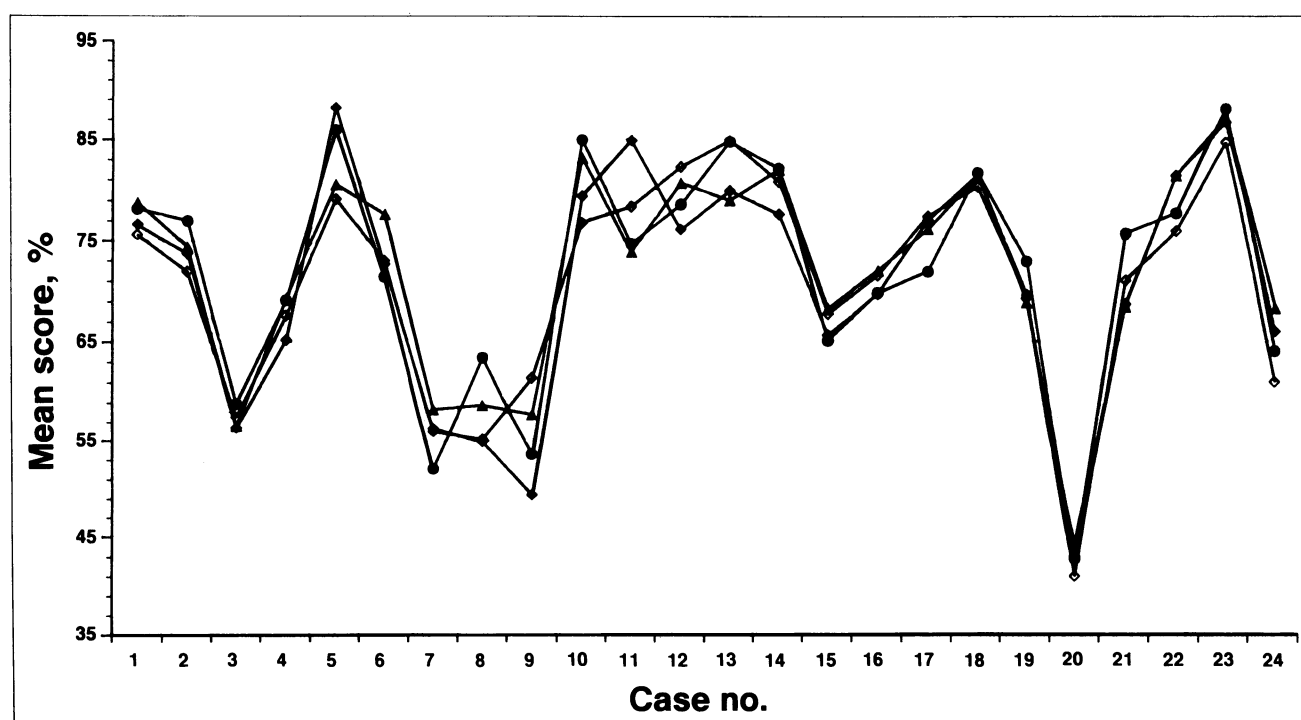


Fig. 2: Mean case scores in the spring of 1991 by centre. Triangles represent Université de Montréal, circles represent Université de Sherbrooke, black diamonds represent McGill University, and white diamonds represent Université Laval.

this newly developed examination, because all candidates had already successfully completed their residency training program. This approach is in agreement with what is usually reported in the literature when a pass-fail standard is needed for tests involving standardized patients.⁷ A success rate of 97% was obtained for the spring 1990 exam, 100% for the fall 1990 exam and 95% for the spring 1991 exam (Table 1).

CFPC examination

The success rate of the candidates was 91% for the 1990 and spring 1991 exams. These rates were similar to those obtained by other Canadian candidates and indicate that Quebec family medicine training programs have maintained certification standards despite a threefold increase in the total number of residents in less than 3 years.¹⁶

Licensure decision

The decision on who should be licensed followed simple rules: candidates who passed the OSCE and the CFPC examination were granted a licence. Those who failed both examinations were not. For those who failed only one examination a board of five family physicians decided whether to grant a licence after reviewing the candidate's in-training evaluations and examination performances.

The overall success rate was 92% in 1990 and 92% in the spring of 1991.

Discussion

During the months preceding the first session of the OSCE two questions were constantly raised: Is the exam feasible? Will it be credible?

Feasibility

OSCEs have been used worldwide for formative and summative purposes and at all levels of training.^{3,4} Feasibility of a large-scale multicentre OSCE is more difficult to determine. Even though a few examples have been reported in the literature^{7,17,18} the Quebec OSCE is the first one to be used on such a large scale as a licensing examination. Planning, committed leaders and organizers, and collaboration of over 350 people were necessary to administer the exam the first three times.

The question of cost has also been raised. Each candidate pays \$650 to take the OSCE. This fee covers about 60% of the cost; the other 40% is shared by the CPMQ and the four Quebec medical schools, either in money or in the time

spent by teachers in family medicine. The fact that the medical schools view the OSCE as their examination has contributed to maintaining the costs at reasonable levels.

The integration of the CFPC certification exam into the licensing exam required adequate logistics, but the feasibility of the process was not questioned, since similar exams had been successfully put in place in many centres for years.

Credibility

The question of credibility relates to validity and reliability. Validity refers to how well the instrument tests what it intends to measure. Content validity refers to how well the test covers the area of competence under consideration. In the Quebec experience the high number of cases and the fact that those who developed the cases were all family physicians from urban, rural and teaching practices assured the content validity of the OSCE. Other forms of validation, such as construct validity studies, are in progress.

Reliability is a quantitative expression of the reproducibility with which the instrument measures the same event on different occasions with different observers. An adequate number of raters, standardized patients and cases must be included.^{7,19} After consideration of all sources of measurement error related to candidates, cases, raters and standardized patients reliability coefficients of OSCEs have been reported to be between 0.41 and 0.88.^{8,19} In the spring 1990 and 1991 OSCEs, for which 8 people played the same patient role in the 7-minute stations and 16 played the same role in the 14-minute stations, the reliability coefficients were 0.644 and 0.736.

Strengths and limitations

One of the strengths of the OSCE used in Quebec is the fact that it was designed to assess not only the basic clinical skills such as history-taking and physical examination but also the ability of physicians to integrate those skills into the entire clinical reasoning process, including making a diagnosis, investigating and treating a problem and communicating with the patient.

One limitation is the exam's complexity because of the number of cases, candidates, centres and languages. This limitation can be overcome, as shown in our experience. Another limitation, more difficult to overcome in such a large-scale OSCE, is the difficulty of involving young children as standardized patients or finding enough patients with the same physical findings. Innovative approaches have to be found to solve these problems.

Conclusion

The use of an OSCE in licensing examinations is feasible. It should better assure the clinical competence of physicians entering into practice and should meet the public concerns for qualified family physicians.

References

1. *Regulations Respecting the Terms and Conditions for the Issuance of Permits of the Corporation professionnelle des médecins du Québec*, GOQ 1987.II.1849
2. Harden RM, Gleeson FA: Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ* 1979; 13: 41-54
3. Handfield-Jones R, Hollingworth GR: CFPC's certification examination: an overview. *Can Fam Physician* 1990; 36: 2069-2074
4. Hart I, Harden R, Walton H (eds): *Newer Developments in Assessing Clinical Competence*, Can-Heal, Montreal, 1986
5. Hart I, Harden R: *Further Developments in Assessing Clinical Competence*, Can-Heal, Montreal, 1987
6. Stillman PL, Regan MB, Philbin M et al: Results of a survey on the use of standardized patients to teach and evaluate clinical skills. *Acad Med* 1990; 65: 288-307
7. Swanson DB, Stillman PL: Use of standardized patients for teaching and assessing clinical skills. *Eval Health Prof* 1990; 13: 79-103
8. Van der Vleuten CPM, Swanson DB: Assessment of clinical skills with standardized patients: state of the art. *Teach Learn Med* 1990; 2: 58-76
9. Corporation professionnelle des médecins du Québec, Facultés de médecine des universités Laval, McGill, Montréal et Sherbrooke: *Aspects législatifs, déontologiques et organisationnels de la pratique médicale au Québec*, Corporation professionnelle des médecins du Québec, Montreal, 1992: 92
10. Wilkinson L: *Systat: the System for Statistics*, Systat Inc, Evanston, Ill, 1989: 444-464
11. Barlett MS: Multivariate analysis. *J R Stat Soc* 1947; 9 (ser B): 176-197
12. Keselman HJ, Rogan JC: A comparison of the modified Tukey and Scheffé methods of multiple comparisons for pairwise contrasts. *J Am Stat Assoc* 1978; 73: 47-51
13. Kramer CY: Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics* 1956; 12: 307-310
14. Stenson H, Wilkinson L: *Testat: a Supplementary Module for Systat*, Systat Inc, Evanston, Ill, 1986: 2-10
15. Cronbach L: Coefficient alpha and the internal structure of test. *Psychometrika* 1951; 16: 297-334
16. Lescop J, Davis MWL: Quebec's compulsory 2-year training program. *Can Fam Physician* 1990; 36: 2045-2050
17. Petrusa ER, Blackwell TA, Carline J et al: A multi-institutional trial of an objective structured clinical examination. *Teach Learn Med* 1991; 3: 86-94
18. Stillman P, Regan MB, Swanson DB et al: An assessment of the clinical skills of fourth-year students at four New England medical schools. *Acad Med* 1990; 65: 320-326
19. Swanson DB, Norcini JJ: Factors influencing reproducibility of tests using standardized patients. *Teach Learn Med* 1989; 1: 158-166

Conferences

continued from page 1731

May 8-11, 1993: Council of Biology Editors 37th Annual Meeting
Sheraton Harbor Island Hotel, San Diego
Council of Biology Editors, 200-111 East Wacker Dr.,
Chicago, IL 60601-4298; (312) 616-0800,
fax (312) 372-7723

May 18-21, 1993: 5th International Conference on Cerebral Vasospasm
Hotel MacDonald and the University of Alberta,
Edmonton, and Jasper Park Lodge, Jasper, Alta.
Canadian Congress of Neurological Sciences, Ste. 810,
906-12 Ave. SW, Calgary, AB T2R 1K7;
(403) 229-9544; fax (403) 229-1661

May 19-23, 1993: Spring Symposium on Back Pain
Hyatt Regency Buffalo, Buffalo, NY
Dr. Aubrey A. Swartz, executive director, American Back Society, St. Joseph's Professional Center, 401-2647
E 14th St., Oakland, CA 94601; (510) 536-9929

July 29-31, 1993: A Mid-Summer's Back Break
Hilton Hawaiian Village, Waikiki, Hawaii
Dr. Aubrey A. Swartz, executive director, American Back Society, St. Joseph's Professional Center, 401-2647
E 14th St., Oakland, CA 94601; (510) 536-9929

Aug. 23-27, 1993: 3rd International Congress on Amino Acids and Analogues
Hotel Capsis Beach and Bungalows, Aghia Pelaghia, Crete, Greece
Dr. G. Lubec, Department of Paediatrics, University of Vienna, Währinger Gürtel 18, A 1090 Vienna, Austria;
fax 011-43-1-40400-3238

Aug. 29-Sept. 4, 1993: 13th International Congress of EEG and Clinical Neurophysiology (sponsored by the International Federation of Clinical Neurophysiology)
Vancouver
Secretariat, 645-375 Water St., Vancouver, BC V6B 5C6;
(604) 681-5226, fax (604) 681-2503

Sept. 4-10, 1993: 15th World Congress of Neurology (sponsored by the World Federation of Neurology and the Canadian Neurological Society)
Vancouver
Secretariat, 645-375 Water St., Vancouver, BC V6B 5C6;
(604) 681-5226, fax (604) 681-2503

Dec. 8-12, 1993: Fall Symposium on Back Pain
Hyatt Regency Embarcadero, San Francisco, Calif.
Dr. Aubrey A. Swartz, executive director, American Back Society, St. Joseph's Professional Center, 401-2647
E 14th St., Oakland, CA 94601; (510) 536-9929